

## EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	7627	licheniformis	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	ON	2006/06/07 14:25
L2	3950	sporulation	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	ON	2006/06/07 14:25
L3	34297	spore	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	ON	2006/06/07 14:25
L4	35654	I2 or I3	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	ON	2006/06/07 14:25
L5	85	I4 with I1	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	ON	2006/06/07 14:27
L6	176	I4 same I1	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	ON	2006/06/07 14:26
L7	139787	mutant or mutated or mutation	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	ON	2006/06/07 14:27
L8	4	I7 with I5	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	ON	2006/06/07 14:28

# SCORE Search Results Details for Application 10510408 and Search Result us-10-510-408-2.rup.

[Score Home Page](#)
[Retrieve Application List](#)
[SCORE System Overview](#)
[SCORE FAQ](#)
[Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10510408 and Search Result us-10-510-408-2.rup.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 20:25:57 ; Search time 142 Seconds  
(without alignments)  
1126.956 Million cell updates/sec

Title: US-10-510-408-2  
Perfect score: 875  
Sequence: 1 MYSRSKFKIGLLLLIGSLAA.....HVIEAVKSSGSQHHVEDMKT 173

Scoring table: BLOSUM62  
Gapop 10.0 ; Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_7.2:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	875	100.0	173	2	Q65GP4_BACILD Q65gp4 bacillus li
2	486.5	55.6	170	1	BOFC_BACSU O05391 bacillus su
3	334	38.2	158	2	Q5KWQ8_GEOKA Q5kwq8 geobacillus
4	281.5	32.2	176	2	Q730B0_BACCL1 Q730b0 bacillus ce

5	280.5	32.1	175	2	Q4MS28_BACCE	Q4ms28	bacillus	ce
6	279.5	31.9	176	2	Q634C1_BACCZ	Q634c1	bacillus	ce
7	276.5	31.6	176	2	Q3EJ53_BACTI	Q3ej53	bacillus	th
8	276.5	31.6	176	2	Q817W1_BACCR	Q817w1	bacillus	ce
9	276.5	31.6	176	2	Q81LG6_BACAN	Q81lg6	bacillus	an
10	274.5	31.4	176	2	Q6HDA3_BACHK	Q6hda3	bacillus	th
11	232	26.5	178	2	Q9KDJ0_BACHD	Q9kdj0	bacillus	ha
12	167	19.1	172	2	Q5WHR7_BACSK	Q5whr7	bacillus	cl
13	105	12.0	190	2	Q2ZJH5_CALSA	Q2zjh5	caldicellul	
14	99	11.3	682	2	Q33NB2_METHU	Q33nb2	methanospir	
15	98.5	11.3	712	1	TBPP_NEIMB	Q9k0v0	neisseria	m
16	98.5	11.3	712	2	Q9JPI9_NEIME	Q9jpi9	neisseria	m
17	96	11.0	357	2	Q7NNF5_GLOVI	Q7nnf5	gloeobacter	
18	95	10.9	212	2	Q2SS56_MYCCA	Q2ss56	mycoplasma	
19	93.5	10.7	401	2	Q2ZFS4_CALSA	Q2zfs4	caldicellul	
20	92.5	10.6	318	2	Q9ZDI9_RICPR	Q9zdi9	rickettsia	
21	91.5	10.5	263	2	Q5KSU5_FUGRU	Q5ksu5	fugu rubrip	
22	91	10.4	614	2	O51187_BORBU	O51187	borrelia	bu
23	91	10.4	1040	2	Q4EJ23_LISMO	Q4ej23	listeria	mo
24	89.5	10.2	709	2	Q2SFX8_9GAMM	Q2sfx8	hahella	che
25	89	10.2	276	2	Q6F7N6_ACIAD	Q6f7n6	acinetobact	
26	88.5	10.1	318	2	Q68X33_RICTY	Q68x33	rickettsia	
27	87.5	10.0	435	2	Q5NVF7_PONPY	Q5nvf7	pongo pygma	
28	87.5	10.0	704	2	Q40PH8_DESAC	Q40ph8	desulfuromo	
29	87.5	10.0	757	2	Q39Z96_GEOMG	Q39z96	geobacter	m
30	87	9.9	309	2	Q379K2_RHOPA	Q379k2	rhodopseudo	
31	87	9.9	340	2	Q5NNW0_ZYMMO	Q5nnw0	zymomonas	m
32	87	9.9	802	2	Q60LR3_CAEBR	Q60lr3	caenorhabdi	
33	87	9.9	1172	1	SMC2_SCHPO	P41003	schizosacch	
34	86.5	9.9	425	1	TIG_BACAN	Q81lb8	bacillus	an
35	86.5	9.9	425	1	TIG_BACC1	Q72zv3	bacillus	ce
36	86.5	9.9	425	1	TIG_BACCR	Q812q9	bacillus	ce
37	86.5	9.9	425	1	TIG_BACCZ	Q633x1	bacillus	ce
38	86.5	9.9	425	1	TIG_BACHK	Q6hd53	bacillus	th
39	86.5	9.9	425	2	Q3EVC1_BACTI	Q3evcl	bacillus	th
40	86.5	9.9	425	2	Q4MS78_BACCE	Q4ms78	bacillus	ce
41	86.5	9.9	433	2	Q3TWV4_MOUSE	Q3twv4	mus musculu	
42	86.5	9.9	436	2	Q6AOC9_MOUSE	Q6a0c9	mus musculu	
43	86	9.8	507	2	Q8WQ14_OSTOS	Q8wql4	ostertagia	
44	86	9.8	1329	2	Q8YYT7_ANASP	Q8yyt7	anabaena	sp
45	85.5	9.8	322	2	Q3UCW9_MOUSE	Q3ucw9	mus musculu	

## ALIGNMENTS

## RESULT 1

## Q65GP4\_BACLD

ID Q65GP4\_BACLD PRELIMINARY; PRT; 173 AA.

AC Q65GP4; Q62S53;

DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.

DT 25-OCT-2004, sequence version 1.

DT 07-FEB-2006, entry version 13.

DE BofC.

GN Name=bofC; OrderedLocusNames=BL01145, BLi02902;

OS Bacillus licheniformis (strain DSM 13 / ATCC 14580).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI\_TaxID=279010;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RX PubMed=15383718; DOI=10.1159/000079829;

RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,  
 RA Ehrenreich P., Baeumer S., Henne A., Liesegang H., Merkl R.,  
 RA Ehrenreich A., Gottschalk G.;  
 RT "The complete genome sequence of *Bacillus licheniformis* DSM13, an  
 RT organism with great industrial potential.";  
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;  
 RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,  
 RA Tang M., Lopez de Leon A., Xiang H., Gusti V., Clausen I.G.,  
 RA Olsen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.L.,  
 RA Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N.,  
 RA Ehrlich S.D., Berka R.M.;  
 RT "Complete genome sequence of the industrial bacterium *Bacillus*  
 RT *licheniformis* and comparisons with closely related *Bacillus* species.";  
 RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AE017333; AAU41770.1; -; Genomic\_DNA.  
 DR EMBL; CP000002; AAU24407.1; -; Genomic\_DNA.  
 KW Complete proteome.  
 SQ SEQUENCE 173 AA; 19539 MW; 00319BEEFBB2CFAC CRC64;

Query Match 100.0%; Score 875; DB 2; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 6e-66;  
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYRSKFKIGLLIGSLLAALSFHLEALAEKPAKVQIQLEKVYLDGDVGIENKVEAARTL 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 MYRSKFKIGLLIGSLLAALSFHLEALAEKPAKVQIQLEKVYLDGDVGIENKVEAARTL 60  
 QY 61 EDFKAAYKGWQLIDQKKGFI LRKQVDDISPLSKTNGYIGVTEDGVISTFHGRPGILSEP 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 EDFKAAYKGWQLIDQKKGFI LRKQVDDISPLSKTNGYIGVTEDGVISTFHGRPGILSEP 120  
 QY 121 IQSFFQIDIKRLESRMADDLRKGIPYRTKKEFEHVIEAVKSSGSQHHVEDMKT 173  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 IQSFFQIDIKRLESRMADDLRKGIPYRTKKEFEHVIEAVKSSGSQHHVEDMKT 173

## RESULT 2

## BOFC\_BACSU

ID BOFC\_BACSU STANDARD; PRT; 170 AA.  
 AC 005391;  
 DT 30-MAY-2000, integrated into UniProtKB/Swiss-Prot.  
 DT 01-JUL-1997, sequence version 1.  
 DT 07-MAR-2006, entry version 37.  
 DE Protein bofC precursor (Bypass-of-forespore protein C).  
 GN Name=bofC; OrderedLocusNames=BSU27750;  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
 RC STRAIN=168 / PY79;  
 RX MEDLINE=97177783; PubMed=9025289;  
 RA Gomez M., Cutting S.M.;  
 RT "BofC encodes a putative forespore regulator of the *Bacillus subtilis*

RT sigma K checkpoint.";  
 RL Microbiology 143:157-170(1997).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
 RC STRAIN=168;  
 RA Tosato V., Bolotin A., Bértani I., Valentino I., Bruschi C.V.;  
 RT "A 17.8 kb segment in the spoVB-nadC region of the *Bacillus subtilis*  
 RT 168 chromosome: sequencing and ruv operon identification.";  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.-K., Codani J.-J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emmerson P.T.,  
 RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.-F., Itaya M.,  
 RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,  
 RA Klein C., Kobayashi Y., Koetter P., Koningstein G., Krogh S.,  
 RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,  
 RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Mauel C.,  
 RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,  
 RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,  
 RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,  
 RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,  
 RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,  
 RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,  
 RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serron P.,  
 RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
 RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,  
 RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,  
 RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,  
 RA Weitzenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
 RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E.,  
 RA Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Nature 390:249-256(1997).  
 RN [4]  
 RP STRUCTURE BY NMR OF 31-170, AND SUBUNIT.  
 RX PubMed=16049010; DOI=10.1074/jbc.M506910200;  
 RA Patterson H.M., Brannigan J.A., Cutting S.M., Wilson K.S.,  
 RA Wilkinson A.J., Ab E., Diercks T., de Jong R.N., Truffault V.,  
 RA Folkers G.E., Kaptein R.;  
 RT "The structure of bypass of forespore C, an intercompartmental  
 RT signaling factor during sporulation in *Bacillus*.";  
 RL J. Biol. Chem. 280:36214-36220(2005).  
 CC -!- FUNCTION: Inhibits the spoIVB zymogen from undergoing  
 CC autocatalytic activation by an unknown mechanism, and in this way  
 CC plays a role in the sigma-K checkpoint of sporulation.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Forespore intermembrane space.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

```

CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----
DR   EMBL; X93081; CAA63621.1; -; Genomic_DNA.
DR   EMBL; Y15896; CAB75329.1; -; Genomic_DNA.
DR   EMBL; Z99118; CAB14735.1; -; Genomic_DNA.
DR   PIR; A69596; A69596.
DR   PDB; 2BW2; NMR; A=31-170.
DR   GenomeReviews; AL009126_GR; BSU27750.
DR   SubtiList; BG11917; bofC.
DR   BioCyc; BSUB1423:BSU2771-MONOMER; -.
KW   3D-structure; Complete proteome; Signal; Sporulation.
FT   SIGNAL          1          30      Probable.
FT   CHAIN           31         170      Protein bofC.
FT                                     /FTId=PRO_0000020824.
FT   VARIANT         96         96      S -> F (in bofC1).
FT   STRAND          34         48
FT   STRAND          50         52
FT   STRAND          54         65
FT   HELIX           66         72
FT   STRAND          73         73
FT   TURN            74         75
FT   STRAND          76         82
FT   TURN            83         84
FT   STRAND          85         91
FT   STRAND          93         94
FT   TURN            97         97
FT   HELIX           98        100
FT   TURN           101        101
FT   STRAND         105        108
FT   TURN           109        111
FT   STRAND         112        119
FT   STRAND         121        121
FT   TURN           122        123
FT   STRAND         124        124
FT   STRAND         126        127
FT   TURN           135        137
FT   STRAND         138        138
FT   HELIX          140        148
FT   STRAND         151        151
FT   HELIX          155        168
SQ   SEQUENCE      170 AA;  19296 MW;  2786FEE1DBDFD2A2 CRC64;

```

```

Query Match          55.6%; Score 486.5; DB 1; Length 170;
Best Local Similarity 58.8%; Pred. No. 4.7e-33;
Matches 97; Conservative 26; Mismatches 33; Indels 9; Gaps 1;

```

```

QY      6 KFKIGLLLLIGSLLAALSFHLEALAE-----KPAKVQIQLEKVYLDGDVGIENKVEA 56
      :|  ||:| | :| | : | :      :| :| :||| ||||| ||:| |
Db      3 RFSTAYLLLGLCSAAVFLIGAPSRALGAEEVHYEPLQVHVQLEKVYLDGDVSIHKHEK 62

QY     57 ARTLEDFKAAYKGWQLIDQKKGFILFRKQVDDISPLSKTNGYIGVTEDGVISTFHGRPGI 116
      :::|| ||| || |::|||:::||||:||||||| |||||:::|||||||
Db     63 VFMSDDFWAAYAGWTLVEQKKGYVLFRRQMDDISPLSKVNGYIGVSDNGVISTFHGRPEP 122

QY    117 LSEPIQSFFQIDIKRLESRMADDLRKGIPYRTKKEFEHVIEAVKS 161
      ||||| |||||:::|||| | :| ||||:|||| ||| ||| :|:
Db    123 ASEPIQSFFQIDLERLESHMQKNLLKGIPFRTKAEFEDVIEHMKT 167

```

```

RESULT 3
Q5KWQ8_GEOKA

```

ID Q5KWQ8\_GEOKA PRELIMINARY; PRT; 158 AA.  
 AC Q5KWQ8;  
 DT 01-FEB-2005, integrated into UniProtKB/TrEMBL.  
 DT 01-FEB-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 5.  
 DE Bypass-of-forespore protein C (Forespore regulator of the sigma-K  
 DE checkpoint).  
 GN OrderedLocusNames=GK2593;  
 OS Geobacillus kaustophilus.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.  
 OX NCBI\_TaxID=1462;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=HTA426;  
 RX PubMed=15576355; DOI=10.1093/nar/gkh970;  
 RA Takami H., Takaki Y., Chee G.-J., Nishi S., Shimamura S., Suzuki H.,  
 RA Matsui S., Uchiyama I.;  
 RT "Thermoadaptation trait revealed by the genome sequence of  
 RT thermophilic Geobacillus kaustophilus."  
 RL Nucleic Acids Res. 32:6292-6303(2004).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BA000043; BAD76878.1; -; Genomic\_DNA.  
 KW Complete proteome.  
 SQ SEQUENCE 158 AA; 18254 MW; C5CB88B037E88C36 CRC64;

Query Match 38.2%; Score 334; DB 2; Length 158;  
 Best Local Similarity 45.7%; Pred. No. 3.6e-20;  
 Matches 64; Conservative 30; Mismatches 46; Indels 0; Gaps 0;

Qy 18 LAALSFHLEALAEKPAKVQIQLEKVYLDGVDGIENKVEAARTLEDFKAAYKGWQLIDQKK 77  
 :||: : : : | | : | ||: ||||: : | | : : : | :|||:  
 Db 10 IAAIVLPVHSASAPVKMTIVLERQYLDGEMSEEKVTETVDSMTEIWKKYRGWQLVTLDD 69  
 Qy 78 GFILFRKQVDDISPLSKTNGYIGVTEGIVSTFHGRPGILSEPIQSFFQIDIKRLESRMA 137  
 | :||| : :||| | |||| | :| :|| :| | :||| :| | ||||| : :|||  
 Db 70 QTIVFRKTINDISPLLKTNGYFGITDDGTLSIFNGKPGRSSEIIQSFFQIDVQKLESRQQ 129  
 Qy 138 DDLRKGIPYRTKKEFEHVIE 157  
 : | :||| :| : :| |||  
 Db 130 EKLKKGIRVLSKERYEQVIE 149

## RESULT 4

Q730B0\_BACC1  
 ID Q730B0\_BACC1 PRELIMINARY; PRT; 176 AA.  
 AC Q730B0;  
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.  
 DT 05-JUL-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 9.  
 DE Bypass-of-forespore protein C, putative.  
 GN ORFNames=BCE\_4507;  
 OS Bacillus cereus (strain ATCC 10987).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;  
 OC Bacillus cereus group.  
 OX NCBI\_TaxID=222523;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX PubMed=14960714; DOI=10.1093/nar/gkh258;

```

RA   Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA   Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA   Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT   "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT   adaptations and a large plasmid related to Bacillus anthracis pXO1.";
RL   Nucleic Acids Res. 32:977-988(2004).
CC   -----
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----
DR   EMBL; AE017194; AAS43408.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE    176 AA;  20490 MW;  B47F265C88E5D950 CRC64;

Query Match          32.2%;  Score 281.5;  DB 2;  Length 176;
Best Local Similarity 43.7%;  Pred. No. 1.1e-15;
Matches 62;  Conservative 25;  Mismatches 54;  Indels 1;  Gaps 1;

```

RESULT 5

```

ID      Q4MS28_BACCE    PRELIMINARY;    PRT;    175 AA.
AC      Q4MS28;
DT      02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT      02-AUG-2005, sequence version 1.
DT      07-FEB-2006, entry version 2.
DE      BofC protein.
GN      ORFNames=BCE_G9241_4479;
OS      Bacillus cereus G9241.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC      Bacillus cereus group.
OX      NCBI_TaxID=269801;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=G9241;
RX      PubMed=15155910; DOI=10.1073/pnas.0402414101;
RA      Hoffmaster A.R., Ravel J., Rasko D.A., Chapman G.D., Chute M.D.,
RA      Marston C.K., De B.K., Sacchi C.T., Fitzgerald C., Mayer L.W.,
RA      Maiden M.C.J., Priest F.G., Barker M., Jiang L., Cer R.Z.,
RA      Rilstone J., Peterson S.N., Weyant R.S., Galloway D.R., Read T.D.,
RA      Popovic T., Fraser C.M.;
RT      "Identification of anthrax toxin genes in a Bacillus cereus associated
RT      with an illness resembling inhalation anthrax.";
RL      Proc. Natl. Acad. Sci. U.S.A. 101:8449-8454(2004).
CC      -!- CAUTION: The sequence shown here is derived from an
CC      EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC      preliminary data.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms

```



CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; AAEK01000010; EAL14975.1; -; Genomic\_DNA.

SQ SEQUENCE 175 AA; 20211 MW; 4056A6CA492FABE7 CRC64;

Query Match 32.1%; Score 280.5; DB 2; Length 175;

Best Local Similarity 43.4%; Pred. No. 1.4e-15;

Matches 63; Conservative 27; Mismatches 50; Indels 5; Gaps 2;

```

Qy      28 LAEKPAKVQIQLEKVYLDGVDVGIENKVEAARTLEDFKAAYKGWQLIDQKKGFILFRKQVD 87
        : || :| | ||::|:|:| | | | | | | | | | | | | | | | | | | | | | | |
Db      32 VTEKEPEVTILLERMYVDGEVSEEIFTEKVADLEKFLQQYKEWQLVDRDDVQIVLQKKVD 91

Qy      88 DISPLSKTNGYIGVTEDGVISTFHGRPGILSEPIQSFFQIDIKRLESRMADDLRKGIPYR 147
        ||||| ||:| | ||:|:|:| | | | | | | | | | | | | | | | | | | | | |
Db      92 DISPLLKTSYFGVSEEGILQIFKGVVPK-SDNAIHSFFQIDMKKLESYERAKLKRGIK 150

Qy      148 TKKEFEHVIEAVKSSGSQHHVEDMK 172
        :|: | | | :| | | :|: |
Db      151 SKEGFVKTIEKMK----QYAVQNK 171

```

#### RESULT 6

Q634C1\_BACCZ

ID Q634C1\_BACCZ PRELIMINARY; PRT; 176 AA.

AC Q634C1;

DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.

DT 25-OCT-2004, sequence version 1.

DT 07-FEB-2006, entry version 10.

DE Probable bypass-of-forespore protein C.

GN Name=bofC; OrderedLocusNames=BCE33L4167;

OS Bacillus cereus (strain ZK / E33L).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;

OC Bacillus cereus group.

OX NCBI\_TaxID=288681;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,

RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,

RA Richardson P., Rubin E., Tice H.;

RT "Complete genome sequence of Bacillus cereus ZK.";

RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; CP000001; AAU16102.1; -; Genomic\_DNA.

KW Complete proteome.

SQ SEQUENCE 176 AA; 20474 MW; AEAE46A6CA41BE34 CRC64;

Query Match 31.9%; Score 279.5; DB 2; Length 176;

Best Local Similarity 43.4%; Pred. No. 1.7e-15;

Matches 63; Conservative 27; Mismatches 50; Indels 5; Gaps 2;

```

Qy      28 LAEKPAKVQIQLEKVYLDGVDVGIENKVEAARTLEDFKAAYKGWQLIDQKKGFILFRKQVD 87
        : || :| | ||::|:|:| | | | | | | | | | | | | | | | | | | | | | | |
Db      32 VTEKEPEVTILLERMYVDGEVSEEIFTEKVADLEKFLQQYKEWQLVDRDDVQIVLQKKVD 91

Qy      88 DISPLSKTNGYIGVTEDGVISTFHGRPGILSEPIQSFFQIDIKRLESRMADDLRKGIPYR 147
        ||||| ||:| | ||:|:|:| | | | | | | | | | | | | | | | | | | | | |
Db      92 DISPLLKTSYFGVSEEGILQIFKGVVPK-SDNAIHSFFQIDMKKLESYERVKLKRGIK 150

```

QY 148 TKKEFEHVIEAVKSSGSQHHVEDMK 172  
 :|: | || :| |: |:|  
 Db 151 SKEGFVKTIEMK-----QYAVQNKK 171

## RESULT 7

Q3EJ53\_BACTI

ID Q3EJ53\_BACTI PRELIMINARY; PRT; 176 AA.

AC Q3EJ53;

DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.

DT 08-NOV-2005, sequence version 1.

DT 07-FEB-2006, entry version 3.

DE BofC protein.

GN ORFNames=RBTH\_01388;

OS *Bacillus thuringiensis* serovar israelensis ATCC 35646.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;

OC *Bacillus cereus* group.

OX NCBI\_TaxID=339854;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 35646;

RA Anderson I., Sorokin A., Kapatral V., Reznik G., Bhattacharya A.,

RA Mikhailova N., Burd H., Joukov V., Kaznadzey D., Walunas T.,

RA D'Souza M., Larsen N., Pusch G., Liolios K., Grechkin Y., Lapidus A.,

RA Goltsman E., Chu L., Fonstein M., Ehrlich D., Overbeek R.,

RA Kyrpides N., Ivanova N.;

RT "Comparative genome analysis of *Bacillus cereus* group genomes withRT *Bacillus subtilis*.";

RL Submitted (SEP-2005) to the EMBL/GenBank/DDBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

-----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; AAJM01000838; EAO51343.1; -; Genomic DNA.

SQ SEQUENCE 176 AA; 20535 MW; 744FD0FDEF8EB665 CRC64;

Query Match 31.6%; Score 276.5; DB 2; Length 176;

Best Local Similarity 42.8%; Pred. No. 3e-15;

Matches 62; Conservative 28; Mismatches 50; Indels 5; Gaps 2;

QY 28 LAEKPAKVQIQLEKVYLDGVDVGIENKVEAARTLEDFKAAYKGWQLIDQKKGFILFRKQVD 87

: :| :| | ||::|::| | | || | || ||::| : :|::|

Db 32 VTKKEPEVTILLERMYVDGEVSEEILTEKVVDLEKFLQQYKEWQLVDRDDVQIVLQKKVD 91

QY 88 DISPLSKTNGYIGVTEDGVISTFHGRPGILSEPIQSFFQIDIKRLESRMADDLRKGIPYR 147

|||| |::| |::|::| | | | |||||::|::| |::| :

Db 92 DISPLLKTSYFGVSEEGILQIFKGVPK-SDNAIHSFFQIDMKKLESYERAKLKRGIK 150

QY 148 TKKEFEHVIEAVKSSGSQHHVEDMK 172

:|: | || :| |: |:|

Db 151 SKEGFVRTIEKMK-----QYAVQNKK 171

## RESULT 8

Q817W1\_BACCR

ID Q817W1\_BACCR PRELIMINARY; PRT; 176 AA.

AC Q817W1;

DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2003, sequence version 1.  
 DT 07-FEB-2006, entry version 10.  
 DE BofC protein.  
 GN OrderedLocusNames=BC4417; ORFNames=BC\_4417;  
 OS Bacillus cereus (strain ATCC 14579 / DSM 31).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;  
 OC Bacillus cereus group.  
 OX NCBI\_TaxID=226900;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;  
 RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,  
 RA Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,  
 RA Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,  
 RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,  
 RA Overbeek R., Kyrpides N.C.;  
 RT "Genome sequence of Bacillus cereus and comparative analysis with  
 RT Bacillus anthracis."  
 RL Nature 423:87-91(2003).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AE016877; AAP11330.1; -; Genomic\_DNA.  
 DR BioCyc; BCER226900:BC4417-MONOMER; -.  
 KW Complete proteome.  
 SQ SEQUENCE 176 AA; 20442 MW; 7642688DF1C20982 CRC64;  
  
 Query Match 31.6%; Score 276.5; DB 2; Length 176;  
 Best Local Similarity 42.8%; Pred. No. 3e-15;  
 Matches 62; Conservative 27; Mismatches 51; Indels 5; Gaps 2;  
  
 QY 28 LAEKPAKVQIQLEKVYLDGVDVGIENKVEAARTLEDFKAAYKGWQLIDQKKGFILFRKQVD 87  
 : :| :| | ||::|:|:| | | | | | | | | | | :|:|  
 Db 32 VTKKEPEVTILLERMYVDGEVSEIILTEKVTDLKFLQYKEWQLVDRDDVQIVLQKKVD 91  
  
 QY 88 DISPLSKTNGYIGVTEDGVISTFHGRPGILSEPIQSFFQIDIKRLESRMADDLRKGIPYR 147  
 |||| |::| |::|:|:| | | | | | | | | | | :|:| :  
 Db 92 DISPLLKTSYFGVSEEGILQIFKGVPK-SDNAIHSFFQIDMKKLESYERAKLKRGIK 150  
  
 QY 148 TKKEFEHVIEAVKSSGSQHHVEDMK 172  
 :|: | | :| | :| :| :|  
 Db 151 SKEGFVKTIEKMK----QYAVQSKK 171

## RESULT 9

## Q81LG6\_BACAN

ID Q81LG6\_BACAN PRELIMINARY; PRT; 176 AA.  
 AC Q81LG6; Q6HSX3; Q6KM61;  
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2003, sequence version 1.  
 DT 07-FEB-2006, entry version 15.  
 DE Bypass-of-forespore protein C, putative.  
 GN OrderedLocusNames=BA4653, BAS4318, GBAA4653; ORFNames=BA\_4653;  
 OS Bacillus anthracis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;  
 OC Bacillus cereus group.  
 OX NCBI\_TaxID=1392;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Ames / isolate Porton;  
 RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;  
 RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,  
 RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,  
 RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,  
 RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,  
 RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,  
 RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,  
 RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,  
 RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,  
 RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,  
 RA Salzberg S.L., Thomason B., Friedlander A.M., Koehler T.M.,  
 RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;  
 RT "The genome sequence of *Bacillus anthracis* Ames and comparison to  
 RT closely related bacteria.";  
 RL Nature 423:81-86(2003).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Ames ancestor;  
 RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,  
 RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,  
 RA Fraser C.M.;  
 RT "Bacillus anthracis comparative genomics.";  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Sterne;  
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,  
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,  
 RA Richardson P., Rubin E., Tice H.;  
 RT "Complete genome sequence of *Bacillus anthracis* Sterne.";  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AE016879; AAP28356.1; -; Genomic\_DNA.  
 DR EMBL; AE017334; AAT33775.1; -; Genomic\_DNA.  
 DR EMBL; AE017225; AAT56616.1; -; Genomic\_DNA.  
 DR TIGR; BA4653; -.  
 DR TIGR; GBAA4653; -.  
 KW Complete proteome.  
 SQ SEQUENCE 176 AA; 20435 MW; B2DE122C23A63145 CRC64;

Query Match 31.6%; Score 276.5; DB 2; Length 176;  
 Best Local Similarity 42.8%; Pred. No. 3e-15;  
 Matches 62; Conservative 28; Mismatches 50; Indels 5; Gaps 2;

QY 28 LAEKPAKVQIQLEKVYLDGVDVGIENKVEAARTLEDFKAAYKGWQLIDQKKGFILFRKQVD 87  
 : || :| | ||::|::| | | ||: || ||::|: |: :|:|  
 Db 32 VTEKEPEVTILLERMYVDGEVSEEIFTEKVADLENILQQYKEWQLVDRDDVQIVLQKKVD 91  
 QY 88 DISPLSKTNGYIGVTEDGVISTFHGRPGILSEPIQSFFQIDIKRLESRMADDLRKGIPYR 147  
 ||||| ||::| ||::|::| | | | |||||::|:| | :|:| :  
 Db 92 DISPLLKTSYFGVSEEGILQIFKGVPK-SDNAIHSFFQIDMKKLESYERAKLKRGRIRIK 150  
 QY 148 TKKEFEHVIEAVKSSGSQHHVEDMK 172  
 :|: | ||:| |: |:|  
 Db 151 SKEGFVKTIKMK----QYAVQNKK 171

## RESULT 10

Q6HDA3\_BACHK

ID Q6HDA3\_BACHK PRELIMINARY; PRT; 176 AA.

AC Q6HDA3;

DT 19-JUL-2004, integrated into UniProtKB/TrEMBL.

DT 19-JUL-2004, sequence version 1.

DT 07-FEB-2006, entry version 9.

DE Probable bypass-of-forespore protein C.

GN Name=bofC; OrderedLocusNames=BT9727\_4156;

OS *Bacillus thuringiensis* subsp. *konkukian*.OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*;OC *Bacillus cereus* group.

OX NCBI\_TaxID=180856;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=97-27;

RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,

RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,

RA Richardson P., Rubin E., Tice H.;

RT "Complete genome sequence of *Bacillus thuringiensis* 97-27.";

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; AE017355; AAT63714.1; -; Genomic\_DNA.

KW Complete proteome.

SQ SEQUENCE 176 AA; 20412 MW; 88752B059C739EFB CRC64;

Query Match 31.4%; Score 274.5; DB 2; Length 176;

Best Local Similarity 42.8%; Pred. No. 4.4e-15;

Matches 62; Conservative 27; Mismatches 51; Indels 5; Gaps 2;

QY 28 LAEKPAKVQIQLEKVYLDGVDVGIENKVEAARTLEDFKAAYKGWQLIDQKKGFILFRKQVD 87  
 : || :| | |::|::| | | | | |::|: |::| |  
 Db 32 VTEKEPEVTILLERMYVDGEVSEEIFTEKVADLEKILQQYKEWQLVDRDDVQIVLQKKVD 91

QY 88 DISPLSKTNGYIGVTEDGVISTFHGRPGILSEPIQSFFQIDIKRLESRMADDLRKGIPYR 147  
 |||| |::| |::|::| | | | | |::|: |::| |  
 Db 92 DISPLLKTSYFGVSEEGILQIFKGVPK-SDNAIHSFFQIDMKKLESYERAKLRGIRIK 150

QY 148 TKKEFEHVIEAVKSSGSQHHVEDMK 172  
 :|: | | :| |::| |  
 Db 151 SKEGFVKTIKMK----QYAVQNKK 171

## RESULT 11

Q9KDJ0\_BACHD

ID Q9KDJ0\_BACHD PRELIMINARY; PRT; 178 AA.

AC Q9KDJ0;

DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.

DT 01-OCT-2000, sequence version 1.

DT 07-FEB-2006, entry version 16.

DE Forespore regulator of the sigma-K checkpoint.

GN OrderedLocusNames=BH1223;

OS *Bacillus halodurans*.OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.

OX NCBI\_TaxID=86665;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=C-125 / JCM 9153;

RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*  
 RT *halodurans* and genomic sequence comparison with *Bacillus subtilis*."  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BA000004; BAB04942.1; -; Genomic\_DNA.  
 DR PIR; G83802; G83802.  
 DR BioCyc; BHAL86665:BH1223-MONOMER; -.  
 KW Complete proteome.  
 SQ SEQUENCE 178 AA; 20241 MW; 5902515B57EA72C1 CRC64;

Query Match 26.5%; Score 232; DB 2; Length 178;  
 Best Local Similarity 32.7%; Pred. No. 1.8e-11;  
 Matches 49; Conservative 32; Mismatches 65; Indels 4; Gaps 3;

QY 9 IGLLLIGSLLAALSFHLEALAE-KPAKVQIQLEKVVLDGVDGIENKVEAARTLEDFKAAAY 67  
 : |||::| : || : ::||: |::|: | | :| |  
 Db 20 VNFFLISSIVAMGGNQIAVAETNSTSISVRLERHYVDGEISEEYITEPIETIEQLMGQY 79  
 QY 68 KGWQLIDQKKGFIKQVDDISPLSKTNGYIGVTEGVDISTFHGRPGILSEP-IQSFFQ 126  
 || : | :: :| |::|::| ||: :: || :: : | :| ||||:  
 Db 80 HGWIFVSQNGDEMILKKDVPDLSPLTKANGFFSISNDGQLNLYDG--SAEEQPIIQSFFR 137  
 QY 127 IDIKRLESRMADDLRKGIPIYRTKKEFEHVI 156  
 :| |::| : | || | | : : | :  
 Db 138 LDTKKLESLHDQLRSGIRIGTLDDYREVL 167

## RESULT 12

## Q5WHR7\_BACSK

ID Q5WHR7\_BACSK PRELIMINARY; PRT; 172 AA.  
 AC Q5WHR7;  
 DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.  
 DT 23-NOV-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 8.  
 DE Forespore regulator of the sigma-K checkpoint.  
 GN OrderedLocusNames=ABC1553;  
 OS *Bacillus clausii* (strain KSM-K16).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.  
 OX NCBI\_TaxID=66692;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RA Takaki Y., Kageyama Y., Shimamura S., Suzuki H., Nishi S., Hatada Y.,  
 RA Kawai S., Ito S., Horikoshi K.;  
 RT "The complete genome sequence of the alkaliphilic *Bacillus clausii*  
 RT KSM-K16."  
 RL Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AP006627; BAD64088.1; -; Genomic\_DNA.  
 KW Complete proteome.  
 SQ SEQUENCE 172 AA; 19409 MW; 8BBDC809CB1048A2 CRC64;

Query Match 19.1%; Score 167; DB 2; Length 172;  
 Best Local Similarity 26.8%; Pred. No. 5.5e-06;  
 Matches 42; Conservative 34; Mismatches 73; Indels 8; Gaps 3;

```

Qy      7 FKIG----LLLIGSLLAA---LSFHLEALAEKPAKVQIQLEKVYLDGDVGIENKVEAART 59
      | :|      ||:: :      |      :| |      :: | |: ||:: | |
Db      13 FSLGAASFALLLGTFIILGEDASNDSPQVANPPNATEVLRETTFADGNIRTERVEETIWA 72

Qy      60 LEDFKAAYKGWQLIDQKKGFIKQVDDISPLSKTNGYIGVTEDGVISTFHGRPGILSE 119
      ::|| | |: ||::|| | : ||| : |:|| | :||::| | | : | :
Db      73 IDDFWAEYEEWQLVDQSNHFVHFRKTILDLSPEVKEDGYLGIDADRGIVLYTGAQE-AGQ 131

Qy      120 PIQSFFQIDIKRLESRMADDLRKGIPYRTKKEFEHVI 156
      : : :|: :| | :|| : : :::
Db      132 IVTRYSEIEAAKLSESDVAALEQGIRIESLEHLQNL 168
  
```

## RESULT 13

## Q2ZJH5\_CALSA

```

ID   Q2ZJH5_CALSA    PRELIMINARY;   PRT;   190 AA.
AC   Q2ZJH5;
DT   20-DEC-2005, integrated into UniProtKB/TrEMBL.
DT   20-DEC-2005, sequence version 1.
DT   07-FEB-2006, entry version 3.
DE   Hypothetical protein precursor.
GN   ORFNames=CsacDRAFT_1333;
OS   Caldicellulosiruptor saccharolyticus DSM 8903.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
OC   Caldicellulosiruptor.
OX   NCBI_TaxID=351627;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=DSM 8903;
RG   US DOE Joint Genome Institute (JGI-PGF);
RA   Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
RA   Hammon N., Israni S., Pitluck S., Richardson P.;
RT   "Sequencing of the draft genome and assembly of Caldicellulosiruptor
RT   saccharolyticus DSM 8903.";
RL   Submitted (OCT-2005) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=DSM 8903;
RG   US DOE Joint Genome Institute (JGI-ORNL);
RA   Larimer F., Land M.;
RT   "Annotation of the draft genome assembly of Caldicellulosiruptor
RT   saccharolyticus DSM 8903.";
RL   Submitted (NOV-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
CC   -----
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----
DR   EMBL; AALW01000009; EAP43465.1; -; Genomic_DNA.
KW   Hypothetical protein; Signal.
FT   SIGNAL      1      34      Potential.
SQ   SEQUENCE    190 AA;  21788 MW;  C36C0158B923C919 CRC64;
  
```

Query Match 12.0%; Score 105; DB 2; Length 190;  
 Best Local Similarity 22.6%; Pred. No. 1.1;

Matches 38; Conservative 39; Mismatches 75; Indels 16; Gaps 6;

```

Qy      4 RSKFKIGLLLLIGSLLAALSFHLEALAEKPAKVQIQLEKVYLDGDVG--IENKVEAAR--- 58
      :|| |   | | :   :: : : : | : : | | |   ::
Db      32 KSKIKSKQQLKGKNITEIAYIKDSFPQRINEDTILVVRKYFKG-CGHIIEEKSNISKEFV 90

Qy      59 --TLEDFKAAYKGWQLIDQKKGFILFRKQVDDISPLSKTNGYIGVTEDGVISTFHGRPGI 116
      | |||| : : ||: :   :: : :   :| :|   :| : :| :| :
Db      91 NMTKEDFKSLFSGWEIDAFNSKYVVISRTFEGYC----SNHFIISIKDDRVAIFYSQP-V 145

Qy      117 LSEPIQSFFQIDIKRLESRMADDLRKGIPYRTKKEFEHVIEAVKSSGS 164
      : : :   | | | : : ||: |||   || | : | : ||
Db      146 DGDNLKLVTPISIDDLPEKEVEDLKKGI---VVNSFEDAIVKIVEDFGS 190

```

## RESULT 14

## Q33NB2\_METHU

```

ID   Q33NB2_METHU   PRELIMINARY;   PRT;   682 AA.
AC   Q33NB2;
DT   06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT   06-DEC-2005, sequence version 1.
DT   07-FEB-2006, entry version 3.
DE   Hypothetical protein.
GN   ORENames=MhunDRAFT_3177;
OS   Methanospirillum hungatei JF-1.
OC   Archaea; Euryarchaeota; Methanomicrobia; Methanomicrobiales;
OC   Methanospirillaceae; Methanospirillum.
OX   NCBI_TaxID=323259;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=JF-1;
RG   US DOE Joint Genome Institute (JGI-PGF);
RA   Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
RA   Hammon N., Israni S., Pitluck S., Richardson P.;
RT   "Sequencing of the draft genome and assembly of Methanospirillum
RT   hungatei JF-1.";
RL   Submitted (OCT-2005) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=JF-1;
RG   US DOE Joint Genome Institute (JGI-ORNL);
RA   Larimer F., Land M.;
RT   "Annotation of the draft genome assembly of Methanospirillum hungatei
RT   JF-1.";
RL   Submitted (NOV-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC   EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC   preliminary data.
CC   -----
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----
DR   EMBL; AALU01000001; EAP17254.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   682 AA;  79233 MW;  DD76EF8BBDE3B67D CRC64;

```

Query Match 11.3%; Score 99; DB 2; Length 682;  
 Best Local Similarity 25.6%; Pred. No. 15;  
 Matches 41; Conservative 32; Mismatches 55; Indels 32; Gaps 6;

Qy 11 LLLIGSLLAALSFHLEALAEKPAKVQIQLEKVYLDGDVG IENKVEAARTLEDFKAAYKGW 70



```

      | | | :|| :| :: || |::|| : | |:: |:: :|| | |
Db      90 LTRISELEKSLSEYLSSIQEKEAQIQILTDLVK-----EHENTASQKIEDFSARIK-- 140

QY      71 QLIDQ-KKGFILFRKQVDDISPLSKTNGYIGVTEGVIDTFHGRPGILSEPIQSFFQIDI 129
      || || : :|| | : | : | : | | | : | :
Db      141 QLTDQLESQKVLFEKEKSDLR--SEYEEILAKTRDEVVRKEHELRTLASD----- 188

QY      130 KRLESRMADDLR-----KGIPYRTKKEFEHVIEAVKS 161
      |::|::|:: | :| : | : :|::|
Db      189 --LKNRIAEEKRVQKDRIREQGEYTKRISELQSFLESARS 226

```

## RESULT 15

## TBPB\_NEIMB

```

ID      TBPB_NEIMB      STANDARD;      PRT;      712 AA.
AC      Q9K0V0;
DT      11-JAN-2001, integrated into UniProtKB/Swiss-Prot.
DT      01-OCT-2000, sequence version 1.
DT      07-MAR-2006, entry version 31.
DE      Transferrin-binding protein 2 precursor (TBP-2).
GN      Name=tbpB; Synonyms=tbp2; OrderedLocusNames=NMB0460;
OS      Neisseria meningitidis serogroup B.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC      Neisseriaceae; Neisseria.
OX      NCBI_TaxID=491;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=MC58 / Serogroup B;
RX      MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
RA      Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA      Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA      Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA      Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA      Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA      Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA      Qin H., Vamathevan J.J., Gill J., Scarlato V., Massignani V., Pizza M.,
RA      Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA      Venter J.C.;
RT      "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT      MC58.";
RL      Science 287:1809-1815(2000).
CC      -!- FUNCTION: Acts as a transferrin receptor and is required for
CC      transferrin utilization (By similarity).
CC      -!- SUBCELLULAR LOCATION: Outer membrane; lipid-anchor (Probable).
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; AE002098; AAF40897.1; -; Genomic_DNA.
DR      PIR; E81196; E81196.
DR      GenomeReviews; AE002098_GR; NMB0460.
DR      TIGR; NMB0460; -.
DR      BioCyc; NMEN491:NMB0460-MONOMER; -.
DR      InterPro; IPR000437; Prok_lipoprot_S.
DR      InterPro; IPR001677; Transferrin_bd.
DR      Pfam; PF01298; Lipoprotein_5; 1.
DR      PROSITE; PS00013; PROKAR_LIPOPROTEIN; FALSE_NEG.
KW      Complete proteome; Lipoprotein; Membrane; Outer membrane; Palmitate;
KW      Receptor; Signal.
FT      SIGNAL      1      20      By similarity.
FT      CHAIN      21      712      Transferrin-binding protein 2.

```

FT /FTId=PRO\_0000018193.  
 FT LIPID 21 21 N-palmitoyl cysteine (Probable).  
 FT LIPID 21 21 S-diacylglycerol cysteine (Probable).  
 SQ SEQUENCE 712 AA; 77417 MW; 77EA248941E8EF0C CRC64;

Query Match 11.3%; Score 98.5; DB 1; Length 712;  
 Best Local Similarity 25.5%; Pred. No. 17;  
 Matches 40; Conservative 21; Mismatches 55; Indels 41; Gaps 7;

Qy 30 EKPAKVQIQLEKVYLDGD-----VGIENKVEAARTLEDFKAAYKG 69  
 | | : : : || | | | | | | : | : || | |  
 Db 99 ELPKRQKSVIEKVETDSDNNIYSSPYLKPSNHQNGNTGNGINQPKNQAKDYENFKYVYSG 158

Qy 70 WQLIDQKKGFILFRKQVDDISPLSKTNGYIGVTEDGVISTFHGRPGILSEPIQSFFQIDI 129  
 | : | | : | | | : || | : || : | : : : |  
 Db 159 WFYKHAKREFNL-----KVEPKSAKNG-----DDGYI-FYHGKEP--SRQLPASGKITY 204

Qy 130 KRLESRMADDLRKGIPYRTKKEFEHVIEAVKSSGSQH 166  
 | : | | : || : | : | : | | : :  
 Db 205 KGV-WHFATDTKKG-----QKFREIIQPSKSQGDY 234

Search completed: June 2, 2006, 20:29:51  
 Job time : 146 secs

SCORE 1.3 BuildDate: 12/06/2005

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 20:27:32 ; Search time 22 Seconds  
(without alignments)  
756.614 Million cell updates/sec

Title: US-10-510-408-2  
Perfect score: 875  
Sequence: 1 MYRSKFKIGLLIGSLLAA.....HVIEAVKSSGSQHHVEDMKT 173

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	486.5	55.6	170	2	A69596	forespore sigma-K
2	232	26.5	178	2	G83802	forespore sigma-K
3	98.5	11.3	712	2	E81196	transferrin-bindin
4	92.5	10.6	318	1	E71690	minor teichoic aci
5	91	10.4	614	2	E70120	hypothetical prote
6	87	9.9	1172	2	S51623	cut14 protein - fi
7	86	9.8	1329	2	AE1901	WD-repeat containi
8	85.5	9.8	435	2	I49327	clathrin-associate
9	85.5	9.8	435	2	A31596	clathrin coat asse
10	84.5	9.7	435	2	G02088	assembly protein 5
11	84	9.6	338	1	DENDG	glyceraldehyde-3-p
12	83.5	9.5	180	2	C89929	conserved hypothet
13	83	9.5	328	2	F83599	hypothetical prote
14	82	9.4	1281	1	GNMSLL	retrovirus-related
15	82	9.4	1375	2	T37672	probable DNA repai
16	81.5	9.3	242	2	A89852	conserved hypothet
17	81.5	9.3	435	2	JC6563	clathrin-associate
18	81.5	9.3	628	2	H89917	conserved hypothet
19	80.5	9.2	620	2	C69278	chorismate mutase/

20	80.5	9.2	1517	2	F85985	glutamate synthase
21	80	9.1	166	2	D95203	isochorismatase fa
22	80	9.1	328	2	B38477	hypothetical prote
23	80	9.1	370	2	A25004	keratin, 53K type
24	80	9.1	416	2	T52497	hypothetical prote
25	79.5	9.1	553	2	F96982	fision threonyl-tR
26	79.5	9.1	577	2	H72368	proline-tRNA ligas
27	79.5	9.1	1517	1	F65112	glutamate synthase
28	79.5	9.1	1538	2	AF0432	glutamate synthase
29	79	9.0	328	2	B72850	baculovirus repeat
30	79	9.0	348	2	S38148	hypothetical prote
31	79	9.0	353	2	A72410	chorismate mutase/
32	79	9.0	599	2	T48039	hypothetical prote
33	79	9.0	934	2	H88391	protein R06B10.2 [
34	78.5	9.0	295	2	T07730	glyceraldehyde-3-p
35	78.5	9.0	561	2	B90414	hypothetical prote
36	78.5	9.0	722	2	A82617	glycyl-tRNA synthe
37	78.5	9.0	1486	2	AI0906	glutamate synthase
38	78	8.9	338	2	T06781	glyceraldehyde-3-p
39	78	8.9	1291	2	T06692	hypothetical prote
40	77.5	8.9	297	2	E86828	hypothetical prote
41	77.5	8.9	390	2	S25787	hypothetical prote
42	77	8.8	259	2	H64211	guanylate kinase (
43	77	8.8	281	2	F69323	ATP phosphoribosyl
44	77	8.8	390	2	D82922	conserved hypothet
45	77	8.8	461	2	B82155	deoxyribodipyrimid

## ALIGNMENTS

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 20:27:43 ; Search time 183 Seconds  
(without alignments)  
437.903 Million cell updates/sec

Title: US-10-510-408-2  
Perfect score: 875  
Sequence: 1 MYSRSKFKIGLLLLIGSLAA.....HVIEAVKSSGSQHHVEDMKT 173

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA\_Main:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	875	100.0	173	5	US-10-510-408-2	Sequence 2, Appli
2	875	100.0	173	5	US-10-510-386-124	Sequence 124, App
3	98.5	11.3	712	4	US-10-240-218-8	Sequence 8, Appli
4	98.5	11.3	712	4	US-10-320-800-73	Sequence 73, Appl
5	98.5	11.3	712	4	US-10-220-481-85	Sequence 85, Appl
6	91	10.4	592	5	US-10-994-726-346	Sequence 346, App
7	91	10.4	612	5	US-10-994-726-345	Sequence 345, App
8	89.5	10.2	435	4	US-10-408-765A-1973	Sequence 1973, Ap
9	88.5	10.1	845	4	US-10-437-963-186548	Sequence 186548,
10	87.5	10.0	431	4	US-10-408-765A-2629	Sequence 2629, Ap
11	87	9.9	1172	4	US-10-369-493-2552	Sequence 2552, Ap
12	86	9.8	1533	4	US-10-437-963-128345	Sequence 128345,
13	85.5	9.8	435	5	US-10-367-057-61	Sequence 61, Appl
14	84	9.6	338	6	US-11-188-298-2895	Sequence 2895, Ap
15	84	9.6	980	4	US-10-437-963-128383	Sequence 128383,
16	82.5	9.4	256	4	US-10-424-599-167485	Sequence 167485,

17	82	9.4	1375	5	US-10-732-923-8606	Sequence 8606, Ap
18	81.5	9.3	618	3	US-09-815-242-10864	Sequence 10864, A
19	81.5	9.3	628	3	US-09-815-242-12616	Sequence 12616, A
20	81.5	9.3	628	4	US-10-282-122A-44243	Sequence 44243, A
21	81.5	9.3	640	3	US-09-815-242-5296	Sequence 5296, Ap
22	81.5	9.3	770	4	US-10-282-122A-57088	Sequence 57088, A
23	80	9.1	166	4	US-10-282-122A-73643	Sequence 73643, A
24	80	9.1	166	5	US-10-472-928-3600	Sequence 3600, Ap
25	80	9.1	166	5	US-10-617-320-3096	Sequence 3096, Ap
26	79.5	9.1	610	5	US-10-450-763-36711	Sequence 36711, A
27	79.5	9.1	906	5	US-10-450-763-48680	Sequence 48680, A
28	79.5	9.1	1449	5	US-10-450-763-60431	Sequence 60431, A
29	79.5	9.1	1517	3	US-09-815-242-10325	Sequence 10325, A
30	79.5	9.1	1517	4	US-10-282-122A-56712	Sequence 56712, A
31	79.5	9.1	1538	4	US-10-282-122A-78229	Sequence 78229, A
32	79	9.0	2273	4	US-10-425-115-314032	Sequence 314032,
33	78.5	9.0	278	5	US-10-450-763-58916	Sequence 58916, A
34	78.5	9.0	295	6	US-11-188-298-12304	Sequence 12304, A
35	78.5	9.0	1486	3	US-09-815-242-13728	Sequence 13728, A
36	78.5	9.0	1486	4	US-10-282-122A-75057	Sequence 75057, A
37	78.5	9.0	1486	4	US-10-282-122A-75654	Sequence 75654, A
38	78	8.9	338	6	US-11-188-298-1497	Sequence 1497, Ap
39	78	8.9	495	4	US-10-424-599-232793	Sequence 232793,
40	77.5	8.9	322	4	US-10-369-493-13946	Sequence 13946, A
41	77.5	8.9	434	4	US-10-425-114-54183	Sequence 54183, A
42	77.5	8.9	434	4	US-10-425-115-291120	Sequence 291120,
43	77.5	8.9	536	3	US-09-738-626-6061	Sequence 6061, Ap
44	77.5	8.9	876	4	US-10-437-963-168060	Sequence 168060,
45	77.5	8.9	1639	4	US-10-437-963-128536	Sequence 128536,

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 20:28:58 ; Search time 14 Seconds  
(without alignments)  
142.912 Million cell updates/sec

Title: US-10-510-408-2  
Perfect score: 875  
Sequence: 1 MYSRSKFKIGLLLLIGSLAA.....HVIEAVKSSGSQHHVEDMKT 173

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA\_New:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	% Query		DB	ID	Description
	Score	Match Length			
1	77	8.8	426	6	US-10-953-349-16804
2	77	8.8	431	6	US-10-953-349-16803
3	75.5	8.6	887	6	US-10-527-411-161
4	74.5	8.5	502	6	US-10-953-349-3682
5	74	8.5	312	6	US-10-953-349-23648
6	74	8.5	641	7	US-11-249-111-76
7	73.5	8.4	348	6	US-10-953-349-6565
8	73.5	8.4	382	6	US-10-953-349-6564
9	73.5	8.4	386	6	US-10-953-349-6563
10	73	8.3	294	6	US-10-953-349-8740
11	73	8.3	338	6	US-10-953-349-8739
12	73	8.3	364	6	US-10-953-349-8738
13	72.5	8.3	259	6	US-10-953-349-16805
14	72.5	8.3	2871	6	US-10-505-928-100

15	71.5	8.2	296	7	US-11-246-957-15	Sequence 15, Appl
16	71.5	8.2	1013	6	US-10-527-411-18	Sequence 18, Appl
17	71	8.1	311	6	US-10-953-349-21965	Sequence 21965, A
18	71	8.1	321	6	US-10-953-349-21964	Sequence 21964, A
19	71	8.1	382	6	US-10-953-349-3683	Sequence 3683, Ap
20	71	8.1	708	7	US-11-293-697-4329	Sequence 4329, Ap
21	70	8.0	291	6	US-10-953-349-8115	Sequence 8115, Ap
22	70	8.0	294	6	US-10-953-349-8114	Sequence 8114, Ap
23	70	8.0	338	6	US-10-953-349-8113	Sequence 8113, Ap
24	70	8.0	787	7	US-11-293-697-3773	Sequence 3773, Ap
25	69.5	7.9	398	7	US-11-219-635-18	Sequence 18, Appl
26	69.5	7.9	673	7	US-11-219-635-2	Sequence 2, Appli
27	69.5	7.9	770	7	US-11-219-635-26	Sequence 26, Appl
28	69	7.9	323	6	US-10-953-349-11054	Sequence 11054, A
29	69	7.9	341	6	US-10-953-349-11053	Sequence 11053, A
30	69	7.9	471	6	US-10-953-349-11052	Sequence 11052, A
31	68.5	7.8	464	6	US-10-953-349-4418	Sequence 4418, Ap
32	68.5	7.8	469	6	US-10-953-349-4417	Sequence 4417, Ap
33	68.5	7.8	504	6	US-10-953-349-4416	Sequence 4416, Ap
34	68	7.8	211	6	US-10-953-349-4200	Sequence 4200, Ap
35	68	7.8	246	6	US-10-953-349-4199	Sequence 4199, Ap
36	68	7.8	272	6	US-10-953-349-4198	Sequence 4198, Ap
37	68	7.8	442	6	US-10-505-928-525	Sequence 525, App
38	67.5	7.7	879	6	US-10-527-411-159	Sequence 159, App
39	67	7.7	296	7	US-11-246-957-20	Sequence 20, Appl
40	67	7.7	863	6	US-10-505-928-158	Sequence 158, App
41	66.5	7.6	216	6	US-10-953-349-28992	Sequence 28992, A
42	66.5	7.6	305	6	US-10-953-349-23286	Sequence 23286, A
43	66	7.5	1085	6	US-10-505-928-175	Sequence 175, App
44	65.5	7.5	419	6	US-10-953-349-21528	Sequence 21528, A
45	65.5	7.5	467	6	US-10-953-349-21527	Sequence 21527, A

## ALIGNMENTS



# SCORE Search Results Details for Application 10510408 and Search Result us-10-510-408-2

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ Comments / Sugg](#)

This page gives you Search Results detail for the Application 10510408 and Search Result us-10-510-408-2.

[start](#)

[Go Back to p](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 20:27:41 ; Search time 49 Seconds  
(without alignments)  
309.037 Million cell updates/sec

Title: US-10-510-408-2  
Perfect score: 875  
Sequence: 1 MYSRSKFKIGLLLLIGSLAA.....HVIEAVKSSGSQHHVEDMKT 173

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5\_COMB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6\_COMB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7\_COMB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/H\_COMB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE\_COMB.pep:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	94	10.7	298	US-09-328-352-6064	Sequence 6064, Ap
2	91	10.4	592	US-09-830-230A-346	Sequence 346, App

3	91	10.4	612	2	US-09-830-230A-345	Sequence 345, App
4	85.5	9.8	435	2	US-09-538-092-963	Sequence 963, App
5	84.5	9.7	435	2	US-09-664-958-12	Sequence 12, Appl
6	83	9.5	354	2	US-09-252-991A-27916	Sequence 27916, A
7	80	9.1	166	2	US-09-583-110-3094	Sequence 3094, Ap
8	80	9.1	166	2	US-09-107-433-3096	Sequence 3096, Ap
9	79	9.0	348	2	US-09-315-794-42	Sequence 42, Appl
10	79	9.0	348	2	US-09-389-341-42	Sequence 42, Appl
11	77.5	8.9	510	2	US-09-605-703B-1144	Sequence 1144, Ap
12	77.5	8.9	534	2	US-09-605-703B-1142	Sequence 1142, Ap
13	76.5	8.7	155	2	US-09-732-210-1631	Sequence 1631, Ap
14	76.5	8.7	463	2	US-09-134-000C-3598	Sequence 3598, Ap
15	76.5	8.7	732	5	PCT-US95-17026-2	Sequence 2, Appli
16	76	8.7	222	2	US-09-248-796A-14267	Sequence 14267, A
17	76	8.7	2954	2	US-09-150-867-1	Sequence 1, Appli
18	76	8.7	2954	3	US-09-724-584-1	Sequence 1, Appli
19	75.5	8.6	341	2	US-09-252-991A-21448	Sequence 21448, A
20	75.5	8.6	463	2	US-08-792-295-1	Sequence 1, Appli
21	75.5	8.6	463	2	US-09-076-432-1	Sequence 1, Appli
22	75.5	8.6	523	2	US-10-104-047-3546	Sequence 3546, Ap
23	75	8.6	2183	2	US-08-746-111-5	Sequence 5, Appli
24	74.5	8.5	343	2	US-09-807-258-24	Sequence 24, Appl
25	74.5	8.5	427	2	US-09-591-447D-2	Sequence 2, Appli
26	74.5	8.5	467	2	US-09-134-001C-3579	Sequence 3579, Ap
27	74	8.5	202	2	US-09-134-001C-3860	Sequence 3860, Ap
28	74	8.5	307	2	US-09-023-905A-36	Sequence 36, Appl
29	74	8.5	359	2	US-09-710-279-2454	Sequence 2454, Ap
30	74	8.5	378	2	US-09-134-001C-4464	Sequence 4464, Ap
31	74	8.5	641	2	US-09-724-623-76	Sequence 76, Appl
32	74	8.5	641	3	US-10-288-930-76	Sequence 76, Appl
33	73	8.3	234	2	US-09-107-532A-6670	Sequence 6670, Ap
34	73	8.3	316	2	US-09-107-532A-4924	Sequence 4924, Ap
35	73	8.3	598	2	US-09-512-563C-8	Sequence 8, Appli
36	73	8.3	598	2	US-09-512-563C-12	Sequence 12, Appl
37	73	8.3	598	2	US-10-060-607-8	Sequence 8, Appli
38	73	8.3	598	2	US-10-060-607-12	Sequence 12, Appl
39	73	8.3	598	2	US-10-104-047-2716	Sequence 2716, Ap
40	73	8.3	653	2	US-09-540-236-3128	Sequence 3128, Ap
41	72.5	8.3	433	2	US-09-809-665A-161	Sequence 161, App
42	72.5	8.3	2871	2	US-09-538-092-936	Sequence 936, App
43	72.5	8.3	3542	2	US-10-087-013-2	Sequence 2, Appli
44	72	8.2	301	2	US-09-107-532A-4934	Sequence 4934, Ap
45	72	8.2	377	2	US-09-489-039A-12546	Sequence 12546, A

## ALIGNMENTS

## RESULT 1

US-09-328-352-6064

; Sequence 6064, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 6064

; LENGTH: 298  
 ; TYPE: PRT  
 ; ORGANISM: Acinetobacter baumannii  
 US-09-328-352-6064

Query Match 10.7%; Score 94; DB 2; Length 298;  
 Best Local Similarity 22.4%; Pred. No. 0.024;  
 Matches 39; Conservative 28; Mismatches 55; Indels 52; Gaps 6;

Qy 34 KVQIQLEKQVYLDGVDVGIENKVEAARTLEDFKAAAYKGWQLIDQKKGFILFRKQVDD----I 89  
 |: :|:|: :|| | :| :|: : :|| : :| : | |  
 Db 50 KLDLQVEEKGIDGPVTRVDRYLEQLTIDTLRKSYPKNSFLGEEFGLQEGKGHDADWCWVI 109  
 Qy 90 SPLSKTNGYI-----GVTEGVI-----STFHGRPGILSEPIQSFF 125  
 || | :| |||: ||| | || :||  
 Db 110 DPLDGTQNFINGFPFCISIAVQHKGVTHQGVYIDPVRDELFSASRGRGAVMNQ----- 163  
 Qy 126 QIDIKRLESRMADDLRK-----GIPYRTKKEFEHV-----IEAVKSSGSQ 165  
 :|: :| | | ||| |: | | : || :|:|  
 Db 164 ----RRIRVNVKDSLENTFLAVGHPYRAKRAGEIVSYAEQHFASLLAVTQAGAQ 213

## RESULT 2

US-09-830-230A-346

; Sequence 346, Application US/09830230A  
 ; Patent No. 6902893  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; TITLE OF INVENTION: Lyme Disease Vaccines  
 ; FILE REFERENCE: PB481US  
 ; CURRENT APPLICATION NUMBER: US/09/830,230A  
 ; CURRENT FILING DATE: 2001-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US98/12718  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/057,483  
 ; PRIOR FILING DATE: 1997-09-03  
 ; PRIOR APPLICATION NUMBER: 60/053,344  
 ; PRIOR FILING DATE: 1997-07-22  
 ; PRIOR APPLICATION NUMBER: 60/053,377  
 ; PRIOR FILING DATE: 1997-07-22  
 ; PRIOR APPLICATION NUMBER: 60/050,359  
 ; PRIOR FILING DATE: 1997-06-20  
 ; NUMBER OF SEQ ID NOS: 756  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 346  
 ; LENGTH: 592  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-830-230A-346

Query Match 10.4%; Score 91; DB 2; Length 592;  
 Best Local Similarity 28.5%; Pred. No. 0.15;  
 Matches 57; Conservative 16; Mismatches 83; Indels 44; Gaps 9;

Qy 2 YSRSKFK---IGLLLLIG---SLLAALSF-HLEALAEKPAKVQIQLEKQVYLDGVDVGIENKV 54  
 :|: :| || :| |: |:| :| | || | ||  
 Db 105 FSQGDYKRIAIGTAIHGIYLSVNGAISFKNLNRLIP-----QIYLGAGYYDIISAIEFSK 159  
 Qy 55 EAARTLEDFKAAAYKGWQLIDQKKGFIL-----FRKQVDDISPLSKTNGYIGVTEGVI 109  
 | | | || ||| |:|: | || | :|  
 Db 160 EETNNLYFSSGVYGDIFLISQSGFIKKISFPFKQIIRILDLSKN-----VEKILVRT 214

```

Qy      110 FHGR-----PGILSEPIQSFFQIDIKRLESRMADDLRKGIPY-----RTKKEF 152
      :      | || | ||:      ||      || |      | ||
Db      215 YDNHFYSYINGQWVFIGKLSLQDQDFE-----KSQRMQLAKNKGSIIYLTAYTLRNKKAV 269

Qy      153 EHVEIAVKSSGSQHHVEDMK 172
      :      : :| ||      | | |
Db      270 DERFKFIKDSGMNAVIDFK 289

```

## RESULT 3

US-09-830-230A-345

```

; Sequence 345, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 345
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-830-230A-345

```

```

Query Match          10.4%; Score 91; DB 2; Length 612;
Best Local Similarity 28.5%; Pred. No. 0.15;
Matches 57; Conservative 16; Mismatches 83; Indels 44; Gaps 9;

```

```

Qy      2 YSRSKFK---IGLLLLIG---SLLAALSF-HLEALAEKPAKVQIQLEKVYLDGDVGIENKV 54
      :|: :|  || : | |: |:| :| |  || | | | ||
Db      125 FSQGDYKRIAIGTAIHGIYLSVNGAISFKNLNRLIP-----QIYLGAGYYDIISAIEFSK 179

Qy      55 EAARTLEDFKAAYKGWQLIDQKKGFI-----FRKQVDDISPLSKTNGYIGVTEDGVIST 109
      |  |  |  || || || | :||: | || |  : : |
Db      180 EETNNLYFSSGVYGDIFLISQSGFIKKISFPFKQIIRILDLSKN-----VEKILVRT 234

Qy      110 FHGR-----PGILSEPIQSFFQIDIKRLESRMADDLRKGIPY-----RTKKEF 152
      :      | || | ||:      ||      || |      | ||
Db      235 YDNHFYSYINGQWVFIGKLSLQDQDFE-----KSQRMQLAKNKGSIIYLTAYTLRNKKAV 289

Qy      153 EHVEIAVKSSGSQHHVEDMK 172
      :      : :| ||      | | |
Db      290 DERFKFIKDSGMNAVIDFK 309

```

## RESULT 4

US-09-538-092-963

Query Match 9.8%; Score 85.5; DB 2; Length 435;  
Best Local Similarity 28.2%; Pred. No. 0.4;  
Matches 35; Conservative 20; Mismatches 42; Indels 27; Gaps 6;

Qy	51	ENKVEAARTLE-----DFKAAYKGWQLIDQ--KKGFILFRKQVDDISPLSKTNGYIGVT	102
	:	:       : :   :   : :     :	
Db	70	KQNVNAAMVFEEFLYKMCDVMAAYFG-KISEENIKNNFVLIYELLDEILDF----	GYPQNS 124
Qy	103	EDGVISTFHGRPGILSEPIQSFFQIDIKRLESRMADDL-----RKGIPIYRTKKEFEHV	155
	:	:     :         :   : : :   :       :	
Db	125	ETGALKTFITQQGIKS-----QHQTKEEQSQITSQVTGQIGWRREGIKYRRNELFLDV	177
Qy	156	IEAV	159
	:	:	
Db	178	LESV	181

http://es/ScoreAccessWeb/GetItem.action?AppId=10510408&seqId=530055&ItemName=us-... 6/7/06

US-09-664-958-12

Query Match 9.7%; Score 84.5; DB 2; Length 435;  
 Best Local Similarity 28.2%; Pred. No. 0.52;  
 Matches 35; Conservative 20; Mismatches 42; Indels 27; Gaps 6;

```

Qy      51 ENKVEAARTLE-----DFKAAYKGWQLIDQ--KKGFI LFRKQVDDISPLSKTNGYIGVT 102
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      70 KQNVNAAMVFEFLYKMC DVMAAYFG-KISEENIKNNFLLIYELLDEILDF----GYPQNS 124

Qy      103 EDGVISTFHGRPGILSEPIQSFFQIDIKRLES RMADDL-----RKGI PYRTKKEFEHV 155
      | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db      125 ETGALKTFITQQGIKS-----QHQTKEEQSQITSQVTGQIGWRREGIKYRRNELFLDV 177

Qy      156 IEAV 159
      : | : |
Db      178 LESV 181
  
```

## RESULT 6

US-09-252-991A-27916

; Sequence 27916, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27916

; LENGTH: 354

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27916

Query Match 9.5%; Score 83; DB 2; Length 354;  
 Best Local Similarity 32.1%; Pred. No. 0.57;  
 Matches 26; Conservative 11; Mismatches 28; Indels 16; Gaps 2;

```

Qy      26 EALAEKPAKVQIQLEKVYLDGDVG IENKVEAARTLEDFKAAYKGWQLIDQKKGFI LFRKQ 85
      : | | : | | | : | | | | | | | | | | | | | | | | | | | | | | |
Db      112 KAAAAGMVRAQIQ LAMLYRDGDG GPQDKTEAARWFR--KAAEQGDAAAQNEMGVLYWRGE 169

Qy      86 VDDISPLSKTNGYIGVTEDGV 106
      | | : | |
Db      170 -----GVDQDRV 176
  
```

## RESULT 7

US-09-583-110-3094

; Sequence 3094, Application US/09583110

; Patent No. 6699703

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

```

; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3094
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3094

```

```

Query Match          9.1%; Score 80; DB 2; Length 166;
Best Local Similarity 24.7%; Pred. No. 0.42;
Matches 45; Conservative 27; Mismatches 44; Indels 66; Gaps 11;

```

```

Qy      14 IGSLLAALS FHL EAL EAKPAKVQIQ-----LEKVYLDGDVG IENKVEAARTLEDFKAAAYK 68
      : :| | | :| | | :| | | :| | | | | | | | | |
Db      10 VQNVLVETG FQTKSLLEKISY LQNQARSKNIEIIYVQ---HIEN--SEAQTS ED----- 58

Qy      69 GWQL-----IDQKKGFILFR-----KQ-----VDDISPLS 93
      ||| : ||| :| : | | | | | | | | | |
Db      59 -WQLSALLNRKPAEKVFQKKYNSIFKETGLKEYLDKQGI EKLVL CGMQTEYCVDTSVKVA 117

Qy      94 KTN GYIGVTEDGVISTF HGRPGILSEPIQSFFQIDIKRLES RMADDLRKGIPYRTKKEFE 153
      || : :| :| | | | :| | | | :| | | | | | | | | |
Db      118 FEYGYQLIVPEGAVTTFDG-DDIPAETINEFY E-DI--WEERFADVL-----DYK 163

Qy      154 HV 155
      | :
Db      164 HI 165

```

## RESULT 8

US-09-107-433-3096

; Sequence 3096, Application US/09107433

; Patent No. 6800744

## ; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAG  
; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD/ROM ISO9660

; COMPUTER:

; OPERATING SYSTEM:

; SOFTWARE:

; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3096:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...166
; SEQUENCE DESCRIPTION: SEQ ID NO: 3096:
US-09-107-433-3096

```

Query Match 9.1%; Score 80; DB 2; Length 166;  
Best Local Similarity 24.7%; Pred. No. 0.42;  
Matches 45; Conservative 27; Mismatches 44; Indels 66; Gaps 11;

Qy	14	IGSLLAALS FHLLEALAEKPAKVQIQ-----LEKVYLDGDVGIENKVEAARTLED FKAAYK	68
		: :     : :     : :     : :     : :	
Db	10	VQNVLVETG FQTKSLLEKISYLNQARSKNIEIIYVQ---HIEN--SEAQTSED-----	58
Qy	69	GWQL-----IDQKKGFILFR-----KQ-----VDDISPLS	93
		:     :	
Db	59	-WQLSALLNRKPAEKVFQKKYNSIFKETGLKEYLDKQGIEKLVLCGMQTEYCVDTSVKVA	117
Qy	94	KTNGYIGVTE DGVISTFHGRPGILSEPIQSFFQIDIKRLES RMADDLRKGIPYRTKKEFE	153
		: :   :       :       :       :	
Db	118	FEYGYQLIVPEGAVTTFDG-DDIPAETINEFYE-DI--WEERFADVL-----DYK	163
Qy	154	HV	155
		:	
Db	164	HI	165

```

RESULT 9
US-09-315-794-42
; Sequence 42, Application US/09315794
; Patent No. 6197517
; GENERAL INFORMATION:
; APPLICANT: Roberts, Christopher J.
; TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
; TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION
; TITLE OF INVENTION: DRUGS
; FILE REFERENCE: 9301-053
; CURRENT APPLICATION NUMBER: US/09/315,794

```



; CURRENT FILING DATE: 1999-05-21  
 ; NUMBER OF SEQ ID NOS: 64  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 42  
 ; LENGTH: 348  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-09-315-794-42

Query Match 9.0%; Score 79; DB 2; Length 348;  
 Best Local Similarity 24.1%; Pred. No. 1.6;  
 Matches 42; Conservative 33; Mismatches 67; Indels 32; Gaps 8;

```

Qy      5 SKFKIGLLLLIGSLLAALSFHLEALAEKPAKVQIQLEKVYL-----DGDVGIEN-KVEAAR 58
      |::| ||||| : :| :| ::| :| || : :|| | |
Db      2 SQYKTGLLLIHPAVTTTPELVENTKAQAASKVKFVDQFLINKLNDGSITLENAYETVH 61

Qy      59 TLEDFKAAYKGWQLIDQKKGFILFRKQVDDISPLS---KTNG-YIGVTE----DGVISTF 110
      | : : | | |:: || | : | || ||::: | : : |
Db      62 -----YLTPEAQTDIKFPPKKL--ISVLADSLKPNGSLIGLSDIYKVDALINGF 107

Qy      111 HGRPGILSEPIQSFFQIDIKRLESRMADDLRKGIPYRTKKEFEHVIEAVKSSGS 164
      |::|| : ::| :| : : | :| || : : | : |
Db      108 E----IINEPDYCWIKMDSSKLNQTVSIPLKKKKTNNTKLQSGSKLPTFKKASS 157

```

## RESULT 10

US-09-389-341-42

; Sequence 42, Application US/09389341  
 ; Patent No. 6200803  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Roberts, Christopher J.  
 ; TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL  
 ; TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATIVE  
 ; TITLE OF INVENTION: DRUGS  
 ; FILE REFERENCE: 9301-057  
 ; CURRENT APPLICATION NUMBER: US/09/389,341  
 ; CURRENT FILING DATE: 1999-09-02  
 ; EARLIER APPLICATION NUMBER: 09/315,794  
 ; EARLIER FILING DATE: 1999-05-21  
 ; NUMBER OF SEQ ID NOS: 72  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 42  
 ; LENGTH: 348  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-09-389-341-42

Query Match 9.0%; Score 79; DB 2; Length 348;  
 Best Local Similarity 24.1%; Pred. No. 1.6;  
 Matches 42; Conservative 33; Mismatches 67; Indels 32; Gaps 8;

```

Qy      5 SKFKIGLLLLIGSLLAALSFHLEALAEKPAKVQIQLEKVYL-----DGDVGIEN-KVEAAR 58
      |::| ||||| : :| :| ::| :| || : :|| | |
Db      2 SQYKTGLLLIHPAVTTTPELVENTKAQAASKVKFVDQFLINKLNDGSITLENAYETVH 61

Qy      59 TLEDFKAAYKGWQLIDQKKGFILFRKQVDDISPLS---KTNG-YIGVTE----DGVISTF 110
      | : : | | |:: || | : | || ||::: | : : |
Db      62 -----YLTPEAQTDIKFPPKKL--ISVLADSLKPNGSLIGLSDIYKVDALINGF 107

Qy      111 HGRPGILSEPIQSFFQIDIKRLESRMADDLRKGIPYRTKKEFEHVIEAVKSSGS 164

```

```

      |::||  : ::| :|  :: |::  || :  :  | : |
Db      108 E----IINEPDYCWIKMDSSKLNQTVSIFLKKKKTNNTKLQSGSKLPTFKKASS 157

```

## RESULT 11

US-09-605-703B-1144

; Sequence 1144, Application US/09605703B

; Patent No. 6962989

; GENERAL INFORMATION:

; APPLICANT: Pompejus, Markus

; APPLICANT: Kroger, Burkhard

; APPLICANT: Schroder, Hartwig

; APPLICANT: Zelder, Oskar

; APPLICANT: Haberhauer, Gregor

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL

; TITLE OF INVENTION: PROTEINS

; FILE REFERENCE: BGI-129CP

; CURRENT APPLICATION NUMBER: US/09/605,703B

; CURRENT FILING DATE: 2000-06-27

; PRIOR APPLICATION NUMBER: 60/142,764

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: 60/152,318

; PRIOR FILING DATE: 1999-09-03

; NUMBER OF SEQ ID NOS: 2934

; SEQ ID NO 1144

; LENGTH: 510

; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum

US-09-605-703B-1144

Query Match 8.9%; Score 77.5; DB 2; Length 510;

Best Local Similarity 21.4%; Pred. No. 4.3;

Matches 43; Conservative 29; Mismatches 76; Indels 53; Gaps 9;

```

Qy      12 LLIGSLLAALS FHL EALAEKPAKVQ----IQLEKVYLDGDVG IENKVEAARTLED FKAAY 67
      ||: | |:| |  |:|  :|:  : : :| : |  : : || | |  :
Db      9  LLVAPLTASLVFCNLAVAANA VEVEAESPVVINEVESNSD-PVGDWVELANTDNNNSIDI 67

Qy      68 KGWQLIDQKK-----GFI LFRKQVDDISPLSKTNG---YIGVTEDGVI 107
      || |:| |:  :  : :  | | : | | | |: :| :
Db      68 SGWSLVDDKEDLENALVLPEGTEIESGGYFVIYTDSADYVPTNNTFGGQ EYFGLGKDDTV 127

Qy      108 STFHGRPGILSEPIQSFFQIDI-----KRLES RMADDLRKGIPYRTKKEFEHVIEAV 159
      |  |  | : : : |:  :  |  |  |:|  |  |
Db      128 -TLRNAEG---E VVATYSWKDLGEHAENTYGRIPDMTGDFANTGVPTPGAKNV-----AA 178

Qy      160 KSSGSQ-----HHVE 169
      : || :  |:|
Db      179 EGSGE EEGV VANAQLPFHNVE 199

```

## RESULT 12

US-09-605-703B-1142

; Sequence 1142, Application US/09605703B

; Patent No. 6962989

; GENERAL INFORMATION:

; APPLICANT: Pompejus, Markus

; APPLICANT: Kroger, Burkhard

; APPLICANT: Schroder, Hartwig

; APPLICANT: Zelder, Oskar

; APPLICANT: Haberhauer, Gregor

```

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; TITLE OF INVENTION:  PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE:  2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1142
;   LENGTH: 534
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-605-703B-1142

```

```

Query Match          8.9%;  Score 77.5;  DB 2;  Length 534;
Best Local Similarity 21.4%;  Pred. No. 4.6;
Matches  43;  Conservative  29;  Mismatches  76;  Indels  53;  Gaps  9;

```

```

Qy      12  LLIGSLLAALS FHL EALAEKPAKVQ----IQLEKVYLDGDVG IENKVEAARTLEDFKAA Y 67
      ||: | |:| | |:| : : :| : | : : || | | :
Db      9  LLVAPLTASLVFCNLAVAA NAVEVEAESPVVINEVESNSD-PVG DWVELANTDNNNSIDI 67

Qy      68  KGWQLIDQKK-----GFI LFRKQVDDISPLSKTNG---YIGVTEDGVI 107
      || |:| |:| : : : | | : | | | |: :| :
Db      68  SGWSLVDDKEDLENALVLP EGTEIESGGYFVIYTDSADYVPTNNTFGGQ EYFGLGKDDTV 127

Qy      108 STFHG RPGILSEPIQSFFQIDI-----KRLES RMADDLRKGIPYRTKKEFEH VIEAV 159
      |   |   | : : : | :   |   | :|   |   |
Db      128 -TLRNAEG---EVVATYSWKDLGEHAENTYGRIPDMTGDFANTGVPTPGAKNV-----AA 178

Qy      160 KSSGSQ-----HHVE 169
      : || :   |:||
Db      179 EGSGE EGVVANAQLPFHNVE 199

```

## RESULT 13

US-09-732-210-1631

; Sequence 1631, Application US/09732210

; Patent No. 6573361

## ; GENERAL INFORMATION:

; APPLICANT: Bunkers, Greg J.

; APPLICANT: Liang, Jihong

; APPLICANT: Mittanck, Cindy A.

; APPLICANT: Seale, Jeffrey W.

; APPLICANT: Wu, Yonnie S.

; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use

; FILE REFERENCE: 38-21(15036)B

; CURRENT APPLICATION NUMBER: US/09/732,210

; CURRENT FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 60/169,513

; PRIOR FILING DATE: 1999-12-07

; PRIOR APPLICATION NUMBER: US 60/169,340

; PRIOR FILING DATE: 1999-12-07

; NUMBER OF SEQ ID NOS: 1753

; SEQ ID NO 1631

; LENGTH: 155

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-09-732-210-1631

Query Match 8.7%; Score 76.5; DB 2; Length 155;  
 Best Local Similarity 24.5%; Pred. No. 0.96;  
 Matches 34; Conservative 21; Mismatches 53; Indels 31; Gaps 6;

```

Qy      39 LEKVYLDGDVGIENKVEAARTLEDFKAAAYKGWQLIDQK---KGFI LFRKQVDDISPL--- 92
      : | : | | : | :          | | : | : | | | : | | : : | |
Db      26 INKMMFDGKKSVAEKI-----IYKAFNKIEEKS GEGKIEVFEKALERV RPLVEV 74

Qy      93 -SKTNGYIGVTEDGVISTFHGRPGILSEPIQSFFQIDIKRLES RMADDLR-----K 142
      | : | | | : | | | | : | : | | | | | | | | | |
Db      75 RSRRVG--GATYQVPVEVRASRQQSLS--IRWILEATRKRNERMMVDRLANELMDAASDK 130

Qy      143 GIPYRTKKEFEHVIEAVKS 161
      | : : | : : : | | | :
Db      131 GAAFKKKEDVHKMAEANKA 149

```

## RESULT 14

US-09-134-000C-3598  
 ; Sequence 3598, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 032796-032  
 ; CURRENT APPLICATION NUMBER: US/09/134,000C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6812  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3598  
 ; LENGTH: 463  
 ; TYPE: PRT  
 ; ORGANISM: Enterococcus faecalis  
 ; FEATURE:  
 ; NAME/KEY: MISC\_FEATURE  
 ; LOCATION: (262)..(262)  
 ; OTHER INFORMATION: Amino acid 262 is Xaa wherein Xaa = any amino acid.  
 US-09-134-000C-3598

Query Match 8.7%; Score 76.5; DB 2; Length 463;  
 Best Local Similarity 27.2%; Pred. No. 4.8;  
 Matches 41; Conservative 24; Mismatches 67; Indels 19; Gaps 7;

```

Qy      3 SRSKFKIGLLLLIGSLLAALS FHL EALAEKPAKVQIQLEKVYLDGDVG IENKVEA---ART 59
      | | : | : : | : | : | : | : | : | : | : | : | : |
Db      26 SNRKYIIAVTACPTGIAHTYMAEDALKKKAKEMGV DI-KVETNGSEGIKNRLTAEDIARA 84

Qy      60 LEDFKAAAYKGWQL--IDQKKGFILFRKQVDDISPLSKTNGYIGVTEDGVISTFHGRPGIL 117
      | | | : : | | | | : | | : | | | : | | | |
Db      85 DGVIVAADKKVEMNRFDGKK---LVNRPVSD--GIRKTEELINLAISGEAPT FHGSDSAA 139

Qy      118 SEPIQSFFQIDIKR--LES RMADDLRKGIPY 146
      | : | | : | | : | : | : | : | : | : | : | : |
Db      140 SD-----QEDSAEGSIGSRIYKDL MNGVSH 164

```

## RESULT 15

PCT-US95-17026-2  
; Sequence 2, Application PC/TUS9517026  
; GENERAL INFORMATION:  
; APPLICANT: ZymoGenetics, Inc.  
; APPLICANT: 1201 Eastlake Avenue East  
; APPLICANT: Seattle  
; APPLICANT: WA  
; APPLICANT: USA  
; APPLICANT: 98102  
; TITLE OF INVENTION: Bovine Factor XIII  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 1201 Eastlake Avenue East  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/17026  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, Gary E  
; REGISTRATION NUMBER: 31-648  
; REFERENCE/DOCKET NUMBER: 94-18PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-442-6673  
; TELEFAX: 206-442-6678  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 732 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-17026-2

Query Match 8.7%; Score 76.5; DB 5; Length 732;  
Best Local Similarity 20.6%; Pred. No. 9.5;  
Matches 32; Conservative 25; Mismatches 47; Indels 51; Gaps 6;

```

Qy      32 PAKV-----QIQLEKVYLDGDVGIENKV-----EAARTLEDFKAAY 67
      ||:|          :||: :||: | : :||:          || | | :
Db      332 PARVVTNYFSAHDNDANLQLD-IFLEEDGNVNSKLTKDSVWNYHCWNEAWMTRPDLPVGF 390

Qy      68 KGWQLIDQKKGFI LFRKQVDDISPLSKTNGYIGVTE DGVISTFHGRPGILSEPIQSFFQI 127
      |||::|          :| :||          | : ||          ||
Db      391 GGWQVVDS-----TPQENS DGM YRCGPASVQA I KHG-----HVC FQF 427

Qy      128 DIKRLES RMADDLRKGIPYRTKKEFEHVIEAVKSS 162
      | : : : || : ||: ||: ||: : :
Db      428 DAPFVFAEVNSDL---VYVTAKKD GTHVVEALDTT 459

```

Search completed: June 2, 2006, 20:28:39  
Job time : 50 secs

SCORE 1.3 BuildDate: 12/06/2005

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 20:25:48 ; Search time 85 Seconds  
 (without alignments)  
 930.570 Million cell updates/sec

Title: US-10-510-408-2  
 Perfect score: 875  
 Sequence: 1 MYSRSKFKIGLLLLIGSLAA.....HVIEAVKSSGSQHHVEDMKT 173

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : A\_Geneseq\_8:\*  
 1: geneseqp1980s:\*  
 2: geneseqp1990s:\*  
 3: geneseqp2000s:\*  
 4: geneseqp2001s:\*  
 5: geneseqp2002s:\*  
 6: geneseqp2003as:\*  
 7: geneseqp2003bs:\*  
 8: geneseqp2004s:\*  
 9: geneseqp2005s:\*  
 10: geneseqp2006s:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	875	100.0	173	8	ADG32154	Adg32154 Mutant B_
2	875	100.0	173	8	ADN60330	Adn60330 B. lichen
3	98.5	11.3	712	4	AAE12019	Aae12019 Neisseria
4	98.5	11.3	712	4	AAU27566	Aau27566 Neisseria
5	94	10.7	298	6	ADA34777	Ada34777 Acinetoba
6	92	10.5	547	3	AAG59650	Aag59650 Arabidops
7	92	10.5	593	3	AAG59649	Aag59649 Arabidops
8	92	10.5	638	3	AAG59648	Aag59648 Arabidops
9	91	10.4	592	2	AA119947	Aay19947 B. burgdo
10	91	10.4	612	2	AA119946	Aay19946 B. burgdo
11	89.5	10.2	435	7	ADJ70167	Adj70167 Human hea
12	87.5	10.0	431	7	ADJ70823	Adj70823 Human hea

13	87	9.9	1172	8	ADN19899	Adn19899	Bacterial
14	85.5	9.8	435	7	ADF09560	Adf09560	Human ada
15	84.5	9.7	435	5	ABG34126	Abg34126	Human cla
16	83	9.5	354	7	ABO79170	Abo79170	Pseudomon
17	81.5	9.3	242	6	ABM72224	Abm72224	Staphyloc
18	81.5	9.3	618	4	AAU35271	Aau35271	Enterococ
19	81.5	9.3	628	4	AAU37023	Aau37023	Staphyloc
20	81.5	9.3	628	6	ABU16319	Abu16319	Protein e
21	81.5	9.3	628	6	ABM73039	Abm73039	Staphyloc
22	81.5	9.3	640	4	AAU33800	Aau33800	Staphyloc
23	81.5	9.3	770	6	ABU29164	Abu29164	Protein e
24	80	9.1	166	6	ABU02222	Abu02222	S. pneumo
25	80	9.1	166	6	ABU45719	Abu45719	Protein e
26	80	9.1	166	8	ADK46579	Adk46579	Streptoco
27	80	9.1	166	8	ADR94461	Adr94461	Novel S.
28	80	9.1	166	9	AEA58331	Aea58331	Streptoco
29	80	9.1	343	8	ADN47749	Adn47749	Thermococ
30	79.5	9.1	610	4	ABG06352	Abg06352	Novel hum
31	79.5	9.1	906	4	ABG18321	Abg18321	Novel hum
32	79.5	9.1	1001	9	AED82508	Aed82508	Hyperimmu
33	79.5	9.1	1449	4	ABG30072	Abg30072	Novel hum
34	79.5	9.1	1517	4	AAU34732	Aau34732	E. coli c
35	79.5	9.1	1517	6	ABU28788	Abu28788	Protein e
36	79.5	9.1	1538	6	ABU50305	Abu50305	Protein e
37	79	9.0	348	4	AAB49963	Aab49963	S. cerevi
38	79	9.0	348	6	ABR52821	Abr52821	Protein s
39	79	9.0	348	7	ADK62214	Adk62214	Disease t
40	78.5	9.0	183	6	ABM71074	Abm71074	Staphyloc
41	78.5	9.0	278	4	ABG28557	Abg28557	Novel hum
42	78.5	9.0	1486	4	AAU38135	Aau38135	Salmonell
43	78.5	9.0	1486	6	ABU47730	Abu47730	Protein e
44	78.5	9.0	1486	6	ABU47133	Abu47133	Protein e
45	77.5	8.9	297	5	ABB54975	Abb54975	Lactococc